Ali Bitkilərin Plastid Təyinatlı Nüvə Genləri

H.F. Əliyeva[1], Ə.Ü. Abduləzimova[1], N.Ş. Mustafayev[1], L.M. Süleymanova[2], Z.Ə. Abaszadə, İ.Ə. Şahmuradov[1,2,*]

[1]AMEA Botanika İnstitutu, Badamdar şossesi, 40, Bakı AZ1073, Azərbaycan;
*[2]Azərbaycan Tibb Universiteti, S.Vurğun küç., 167, AZ1022, Azərbaycan; *E-mail: ilham@pgenomics.org*

8 ali bitkidə (birləpəli *Oryza sativa*, *Sorghum bicolor* və *Zea mays*, ikiləpəli *Arabidopsis thaliana*, *Medicago truncatula*, *Populus trichocarpa*, *Glycine max* və *Vitis vinifera*) xloroplast təyinatlı zülal kodlaşdıran (XTZK) nüvə genlərinin axtarışı, həmçinin kəsəkotu və düyünün XTZK nüvə genlərinin potensial promotorlarının hansı sinfə (TATA yaxud qeyri-TATA) mənsub olması baxımından müqayisəli analizi həyata keçirilmişdir. Aşkar olunmuşdur ki, bu bitkilərin hər birinin nüvə genomunda plastid təyinatlı olması güman olan 3000-dən yuxarı zülal (polipeptid) geni kodlaşdırılır. Plastid təyinatlı nüvə genlərinin ümumi sayına görə müxtəlif növlər arasında böyük fərq müşahidə olunur. Bütöv zülal ardıcıllıqlarının 60% və daha yuxarı oxşarlıq dərəcəsi meyarı əsas götürülməklə, müxtəlif tip oxşar gen qrupları müəyyənləşdirilmişdir. 4642 qrupda 3 birləpəli bitkinin hər biri, ən azı, bir gen ilə təmsil olunur. İkiləpəlilərdə 2042 belə qrup aşkar olunmuşdur. *A.thaliana* və *O.sativa* bitkilərinin 29 metabolik yol və ya funksional kompleks ilə bağlı 1159 və 1491 XTZK nüvə geni üçün potensial TATA və/ya qeyri-TATA promotorlar aşkar edilmişdir. Həmin potensial promotorlar içərisində TATA və qeyri-TATA tip promotorların nisbi payı baxımından (arabidopsisdə 41%/59%, düyüdə 26%/74%) qeyri-TATA tip promotorlar üstünlük təşkil edirlər.

Açar sözlər: *Ali bitki, plastid, nüvə, genom, gen, zülal, promotor, kompüter analizi*

GİRİŞ

2000-ci ilin dekabr ayında ali bitkilərin ilk nümayəndəsinin - kəsəkotunun (*Arabidopsis thaliana*) nüvə genomunun (The Arabidopsis Genome Initiative, 2000), 2002-ci ilin aprel ayında isə düyünün (*Oryza sativa* L. *ssp. ındica* və *Oryza sativa* L. *ssp. japonica)* növmüxtəlifliklərinin (Yu et al., 2002; Goff et al., 2002) nüvə genomlarının qaralama variantları oxunmuşdur. Bütövlükdə, son 13 ildə 10-dan çox bitkinin nüvə genomu tam yaxud qaralama variantında, onlarla bitkinin plastid və mitoxondri genomu isə tam oxunmuşdur (http://www.ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html; Hamilton and Buell, 2012). Lakin bu genomların gen tərkibi, bu genlərin funksiyası və ekspressiyasının tənzimlənməsi mexanizmləri ilə bağlı bir çox suallar hələ də cavabsızdır. Həmin suallara cavab tapmaq üçün aparılan tədqiqatlarda geniş tətbiq olunan yanaşmalardan biri təcrübi və bioinformatik üsullarla növdaxili və növlərarası müqayisəli analizdir.

Genlərin genomda təşkili və funksiyası, təkamülü, onların promotor nahiyyələrindəki transkripsiya tənzimləyici elementləri haqqında yeni biliklər əldə olunmasında **ortoloji** (*müxtəlif növlərdə eyni mənşəli və eyni/oxşar funksiyalı*) genlərin müqayisəli analizi ən səmərəli yollardan biridir. Bundan başqa, sistematika və təkamül biologiyası sahələrində aparılan tədqiqatlarda da ortoloq genlərin təsdiqlənmiş dəstlərinin müqayisəli analizi önəmli rol oynayır (Wu et al., 2006; Balaji et al., 2006; Kim et al., 2008). Məsələn, məməli və bitki ortoloq genlərinin promotor nahiyyələrinin konservativ xüsusiyyətlərinin tədqiqi göstərmişdir ki, transkripsiyanın start nöqtələri, TATA-boks və digər mühüm tənzimləyici elementlər onların ətrafındakı digər ardıcıllıqlara nisbətən statistik cəhətdən əhəmiyyətli dərəcədə daha konservativdir. Bu müşahidələr əsasında insan/heyvan və bitki orqanizmlərdə Pol II promotorlarının axtarışı üzrə **PromH** və **PromHP** kompüter proqramları yaradılmışdır və həmin proqramlar, digər müvafiq üsullar ilə müqayisədə, daha yüksək dəqiqliyə (85-95%) malikdirlər (Shahmuradov and Solovyev, 2003; http://www.softberry.com/berry.phtml).

Beləliklə, bu məsələlərin həlli üçün əvvəlcə müvafiq ortoloji genlər kolleksiyası yaradılmalıdır. İndiyə qədər həmin istiqamətdə müəyyən cəhdlər və uğurlar olsa da, bu məsələnin tam həlli çox çətindir.

Ortoloqların müəyyənləşdirilməsində istifadə olunan alqoritmlərin əksəriyyəti iki və daha çox növün gen/zülal yaxud bütöv genom ardıcıllıqlarının müqayisəsinə əsaslanır. Bu müqayisələrdə ən çox istifadə olunan kompüter vasitəsi BLAST proqramıdır (Altschul et al., 1997; http://blast.ncbi.nlm.nih.gov/Blast.cgi).

İndiyə qədər 20 bitki növünü təmsil edən 500000-dən çox məlum və ya güman olunan polipeptid annotasiya olunmuşdur. Bu ardıcıllıqların cüt-cüt (500000 x 500000) BLAST müqayisəsi nəticəsində çoxsaylı potensial ortoloqlar aşkar olunmuşdur. Həmin ortoloqların sonrakı analizi nəticəsində

müxtəlif orqanizmlərdən təxminən 430000 ortoloji ardıcıllıq müəyyən edilmişdir. Bu gün ortoloq və **homoloq** (*eyni bir növə mənsub olan, eyni mənşəli və oxşar*) genlər/zülallar üzrə bir sıra məlumat bazaları vardır, o cümlədən **Homogene** – bir neçə bitki və onurğalı heyvan orqanizmini əhatə edir (http://www.ncbi.nlm.nih.gov/homologene/); **TreeFam** – bir neçə qeyri-heyvan model orqanizm də daxil olsa, əsasən heyvan orqanizmlərini əhatə edir (http://www.treefam.org/); **OrthoMCL** – 150-dən çox müxtəlif genomu əhatə edir, səmərəli axtarış alətlərinə malikdir (http://www.orthomcl.org/orthomcl/); **InParanoid** – genomların cüt-cüt müqayisəsinə əsaslanır və bir neçə genomu əhatə edir (http://inparanoid.sbc.su.se/cgi-bin/index.cgi); **MetaPhOrs** – məlum ortoloq məlumat bazalarının inteqrasiyası əsasında yaradılmışdır (http://orthology.phylomedb.org/?q=home); **Compara** – **Ensembl** məlumat bazasının bir hissəsidir, dəqiq analiz mexanizminə malikdir, lakin interfeysi çətindir (http://www.ensembl.org/info/genome/compara/index.html); **KEGG** – bu məlumat bazasında hüceyrə, orqanizm və ekosistem ilə bağlı metabolizm, molekulyar və digər yönümlü məlumatlar vardır, lakin həmin resurslardan tam məlumatları əldə etmək üçün ödənişli lisenziya tələb olunur (http://www.genome.jp/kegg/).

Genlərin kodlaşdırdıqları zülalların funksiyalarının, həmçinin ortoloji genlərin müəyyənləşdirilməsində ortaya çıxan əsas problemlərdən biri həmin zülalların hüceyrə daxilində və xaricində lokalizasiyasının ("təyinat yerinin") müəyyənləşdirilməsidir.

Ədəbiyyatda *A.thaliana* hüceyrələrində sintez olunan zülalların subhüceyrə lokalizasiyasının mass-spektrometriya (*tandem mass-spektrometriya data analysis, MS/MS*), biokimyəvi təhlil və intakt hüceyrələrdə lüminessent zülal ximerlərinin analizi (*chimeric fusion studies, GFP)* metodları vasitəsi ilə müəyyənləşdirilməsi üzrə təcrübələrin nəticələri əsasında hüceyrənin müxtəlif tərkib hissələrinin (plastid, mitoxondri, peroksisoma, nüvə, plazmatik membran, endoplazmatik retikulum, vakuol) müvafiq surətdə ~2500, ~1800 və ~900 zülalı əhatə edən proteomaları üzrə məlumatlar mövcuddur. Zülalların daşınma yerlərinin, təcrübi araşdırmalarla yanaşı, bioinformatik üsullarla analizi sitoplazmada sintez olunan zülallarını son təyinat ünvanları üzrə biliklərimizin dəqiqləşdirilməsi və genişləndirilməsi üçün ən səmərəli yol hesab olunur (Heazlewood et al., 2005). Zülalların subhüceyrə lokalizasiyasının məhz qeyd olunan inteqrativ şəkildə təhlili nəticəsində kəsəkotu zülallarının hüceyrədaxili təyinat yerləri üzrə SUBA məlumat bazası yaradılmışdır (Heazlewood et al., 2007; Tanz et al., 2013); http://suba3.plantenergy.uwa.edu.au/). Bu bazanın 2014-cü il versiyası subhüceyrə lokalizasiyası GFP və MS/MS metodları vasitəsi ilə müəyyənləşdirilmiş və 1415 elmi məqalədə təqdim olunmuş 10305 müxtəlif zülalı əhatə edir.

Bryant və həmkarları (Bryant et al., 2011) kəsəkotu (*Arabidopsis thaliana*) bitkisi üzrə RIKEN xloroplast funksiyaları məlumat bazasında (http://rarge.psc.riken.jp/chloroplast/) və SeedGenes database məlumat bazasında (www.seedgenes.org) toplanmış mutant variantların və ədəbiyyat məlumatlarının təhlili əsasında xloroplast funksiyalar ilə bağlı 381 nüvə genindən ibarət kolleksiya yaratmışlar: embrion qüsurları (*embryo-defective emb*) törədən119 gen, bitki piqmentləri ilə bağlı 121 gen, 3 qametofit geni və alternativ fenotiplər ilə bağlı 138 gen aşkar olmuşdur ki, embrionların inkişafı amin turşuları, vitaminlər və nukleotidlərin biosintezi pozulanda çox vaxt dayanır, fotosintez prosesində problem yarananda, xlorofil, karatenoidlərin və ya terpenoidlərin sintez səviyyəsi aşağı düşəndə isə davam edir.

Dinq və həmkarları (Ding et al., 2013) Markov zəncirlərinin interpolyasiyası əsasında nüvədə kodlaşan xloroplast genlərini proqnozlaşdırmaq üçün yeni NIM kompüter metodu işləyib hazırlamışdır. Təcrübi məlumatlar əsasında sınaqlar göstərmişdir ki, NIM metodunun orta həssaslığı 92%-dən yuxarı, orta spesifikliyi isə 97%-dən yuxarıdır (bu göstəricilərə görə NIM metodu bu sahədəki ən yaxşı kompüter proqramları sırasındadır).

Qara qovaq (*Populus trichocarpa*) bitkisində fotosintezin tənzimlənməsində iştirak edən genləri bütöv genom miqyasında aşkar etmək və səciyyələndirmək üçün, mikroarrey və BSA (*bulked segregant analysis*) strategiyaları əsasında fotosintez prosesi ilə bağlı ekspressiya olunan 515 nüvə geni (163 up-reqlyasiya və 352 down-reqlyasiya) müəyyən edilmişdir (Wang et al., 2014).

Fotosintetik və bir sıra digər mühüm funksiyaları yerinə yetirən, xloroplast, leykoplast, amiloplast və digər formaları olan plastidlər bitkilərin spesifik və mühüm orqanellalarından biridir. Onların biogenezi və fəaliyyəti plastid və nüvə genomlarının birgə genetik nəzarəti altında baş verir. Müasir ali bitkilərin plastid genomunda təxminən 75 müxtəlif zülal kodlaşdıran gen aşkar edilmişdir (Mache, Lerbs-Mache, 2001). Bu genləri şərti olaraq 4 qrupa daxil etmək olar (Gruissem, Tonkyn, 1993):

(1) genlərin ekspressiyası üçün mühüm olan «ev təsərrüfatı» genləri;

(2) ATF-azadan asılı proteazanın subvahidinin və NADH-dehidrogenaza kompleksinin polipeptidlərinin genləri;

(3) fotosintetik aparatın tərkibinə daxil olan polipeptidlərin genləri;

(4) müxtəlif orqanizmlərdə sayına və tərkib hissəsinə görə variasiya edən hipotetik zülalların genləri (ORF-genlər).

Lakin plastid genomunun rolunu azaltmadan qeyd etmək lazımdır ki, plastid təyinatlı zülalların

mütləq əksəriyyəti (ən azı, 95%-i) nüvə genomunda kodlaşdırılır, sitoplazmada sintez olunur və sonra bu orqanellaya daşınır. Ali bitkilərin nüvə genomları oxunduqdan sonra aparılmış təcrübi və nəzəri araşdırmaların nəticələrinə əsasən, nüvədə müxtəlif metabolik proseslərlə bağlı plastid (və ya bundan sonra sinonim termin kimi işlədilən xloroplast) təyinatlı bir neçə min gen kodlaşdırılır, ancaq bu genlərin tam və dəqiq siyahısı bu gün məlum deyildir (Martin et al., 1998, 2002). Digər tərəfdən, məlumdur ki, yetkin plastidlərin ən yaxşı öyrənilmiş və hazırda da intensiv surətdə tədqiq olunan forması olan xloroplastlardan «verilən» siqnallar da nüvə genlərinin ekspressiyasının tənzimlənməsində müəyyən rol oynayırlar (Pfannschmidt, 2003).

Məlumdur ki, istənilən genin ekspressiyasının ilkin mərhələsi transkripsiyadır. Transkripsiyanın inisiasiyası və tənzimlənməsi əsasən gendən əvvəldə (5‘-nahiyyədə) yerləşən, tərkibində çoxsaylı və müxtəlif tənzimləyici elementlər (TE) olan promotor rayonu və həmin TE-lərlə birləşən tənzimləyici zülal faktorları vasitəsi ilə həyata keçirilir. O cümlədən, proksimal promotorda, yəni genin kodlaşdıran hissəsinə (KH) daha yaxın nahiyyədə yerləşən müəyyən DNT ardıcıllıqları müvafiq RNT polimeraza (zülal kodlaşdıran eukariot genləri üçün RNT polimeraza II) vasitəsi ilə transkripsiyanın start saytını (TSS) müəyyən edirlər.

Hazırda RNT polimeraza II (Pol II) vasitəsi ilə transkripsiya olunan zülal kodlaşdıran genlərin promotorlarının quruluş baxımından 2 əsas sinfi məlumdur: (1) tərkibində genin başlanğıcından təxminən 20-45 nukleotid cütü (nc) əvvəldə yerləşən və TATA-boks adlanan element olan TATA promotorlar; (2) qeyd olunan nahiyyədə TATA-boks olmayan qeyri-TATA promotorlar. Ədəbiyyatdan məlumdur ki, bəzi sinif genlərin (məsələn, ümumi təyinatlı «ev təsərrüfatı genləri»nin) transkripsiyası məhz qeyri-TATA promotorlardan inisiasiya olunur (Novina, Roy, 1996; Pedersen et al., 1999; Smale, 1997; Lemon, Tjian, 2000; Smale, 2001; Ohler at al, 2002; Zhao et al., 2003).

Son illərdə ədəbiyyatda verilən məlumatlara əsasən, xloroplast təyinatlı genlər əsasən qeyri-TATA promotorlarla səciyyələnirlər (Nakamura et al., 2002). Qeyd olunan işdə nüvədə kodlaşan fotosistem I genlərinin promotorlarının arxitekturası, model gen kimi, tütünün *psaDb* genində tədqiq olunmuşdur. Müəyyən edilmişdir ki, bu genin promotor nahiyyəsində TATA-box mövcud deyildir. Bunun əvəzinə bu gendə transkripsiyaının start saytını əhatə edən, primidinlərlə zəngin və «Initiator (İnr)» adlanan element əsas rol oynayır. 232 bitki geninin promotorun tipi baxımından analizi çox maraqlı fakt aşkar etmişdir: fotosintetik genlərdə qeyri-TATA promotorlar mütləq üstünlük təşkil edirlər.

Yuxarıda deyilənləri əsas götürməklə, bu işin əsas məqsədi bir sıra ali bitkilərin xloroplast təyinatlı zülal kodlaşdıran (XTZK) nüvə genlərinin müəyyənləşdirilməsi, həmçinin kəsəkotu və düyünün XTZK nüvə genlərinin potensial promotorlarının hansı sinfə (TATA yaxud qeyri-TATA) mənsub olması baxımından müqayisəli analizi olmuşdur.

MATERİAL VƏ METODLAR

Analiz üçün nüvə genomu tam yaxud qaralama variantında oxunmuş 8 bitkinin - **birləpəli** düyü (*Oryza sativa japonica*; 12 xromosom; NCBI: NC_008394 - Chr1, NC_008395 - Chr2, NC_008396 - Chr3, NC_008397 - Chr4, NC_008398 - Chr5, NC_008399 - Chr6, NC_008400 - Chr7, NC_008401 - Chr8, NC_008402 - Chr9, NC_008403 - Chr10, NC_008404 - Chr11, NC_008405 - Chr12), kalış (*Sorghum bicolor*; 20 xromosom) və qarğıdalı (*Zea mays*; 10 xromosom; http://ftp.maizegdb.org/MaizeGDB/FTP/maize_proteome/proteome.fasta), **ikiləpəli** kəsəkotu (*Arabidopsis thaliana*; 5 xromosom), qara yoncanın (*Medicago truncatula*; 8 xromosom; http://jcvi.org/medicago/index.php), qovaq ağacı (*Populus trichocarpa*; 19 xromosom ftp://ftp.ncbi.nlm.nih.gov/genomes/Populus_trichocarpa/), soya (*Glycine max*; 20 xromosom; ftp://ftp.ncbi.nlm.nih.gov/genomes/Glycine_max/) və şərab üzümünün (*Vitis vinifera*; 19 xromosom; ftp://ftp.ncbi.nlm.nih.gov/genomes/Vitis_vinifera/) müvafiq surətdə 28392, 29448, 63540, 35176, 64123, 28504, 55787 və 21764 məlum yaxud güman olunan zülal ardıcıllıqları, həmçinin kəsəkotu və düyü bitkilərinin 29 müxtəlif metabolik yol və/ya funksional kompleks ilə bağlı, xloroplast təyinatlı zülal kodlaşdıran nüvə genləri (müvafiq surətdə 1779 və 1837 gen) götürülmüşdür (http://chloroplast.net/index.html; cədvəl 1). Həmin genlərin məhz xloroplast təyinatlı olması, yəni onların kodlaşdırdıqları zülalların sintez olunduqdan sonra xloroplasta daşınması aşağıdakı meyarlardan biri və ya bir neçəsi əsasında müəyyənləşdirilmişdir:

(1) xloroplastda lokalizasiyası eksperimental yolla müəyyən edilmiş zülalların genləri;
(2) xloroplastlarda baş verən metabolik proseslərdə iştirakı məlum olan zülalların genləri;
(3) (1) və (2) bəndlərində qeyd olunan zülallara oxşar zülalların genləri.

Seçilmiş gen və zülallar içərisində təkrarlanan ardıcıllıqların olmaması, yəni təkrarsız (hər gendən/zülaldan bir nüsxə olan) kolleksiyaların seçilməsi üçün bütün zülal ardıcıllıqları **BLAST** (Altschul et al., 2004; http://www/ncbi.nlm.nih. gov) və **BLAN** (İ.Şahmuradov; çap olunmamışdır) kom-

Cədvəl 1. Kəsəkotu və düyünün metabolik yollar və funksional komplekslər ilə bağlı, XTZK nüvə genləri qrupları

Ümumi funksional təsnifat qrupu	**Konkret funksional təsnifat qrupu**	**Qrupdakı genlərin sayı**	
		kəsəkotu	**düyü**
1. Genetik sistem	1.1. DNT-nin replikasiyası, reparasiyası və rekombinasiyası	28	30
	1.2. Transkripsiya	35	23
	1.3. Transkriptlərin prosessinqi	33	33
	1.4. Translyasiya	99	113
	1.5. Post-translyasiya modifikasiyası	27	24
2. Fotosintezin işıq reaksiyası	2.1. İşıq reaksiyası	76	57
	2.1. Assemblinq faktorları	8	6
	2.2. Fosforlaşma və defosforlaşma	4	3
	2.3. Birləşdirici faktorlar	10	6
3. Fotosintezin qaranlıq reaksiyaları	3.1. Qaranlıq reaksiyaları	66	74
4. Xloroplast membranı	4.1. Xloroplast membranının trnaslokon sistemləri	28	23
	4.2. Translokatorlar və transportorlar	48	75
	4.3. Şaperonlar və istilik şoku zülalları	28	47
	4.4. Tilakoid membranının translokon sistemləri	11	13
	4.5. Proteazalar	55	46
5. Metabolizm	5.1. Karbohidrat metabolizmi	115	137
	5.2. Amin turşularının sintezi	181	174
	5.3. Azot metabolizmi	148	151
	5.4. Sulfat metabolizmi	53	62
	5.5. Yağ turşuları metabolizmi	121	143
	5.6. piqment sintezi və deqradasiyası	57	52
	5.7. elektron daşıyıcılarının sitezi və deqradasiyası	2	5
	5.7. İşıq reseptorlarının sintezi və deqradasiyası	7	4
	5.8. İkinci dərəcəli metabolizm	271	273
6. İnkişaf	6.1. Xloroplastın biogenezi	40	42
	6.2. Xloroplastın inkişafı və difirensasiyası	43	32
	6.3. Qocalma	4	5
	6.4. Stresə cavab verən genlər	101	93
	6.5. Müdafiə mexanizmi	3	5
CƏMİ:	**29 qrup**	**1770**	**1823**

püter proqramlarının vasitəsi ilə cüt-cüt (növdaxili və növlərarası) müqayisə olunmuş və tam uzunluqlu oxşarlıq dərəcəsi 90%-də yüksək olan ardıcıllıqlardan yalnız biri götürülmüşdür. Bu yolla yuxarıda qeyd olunan 8 növdən müvafiq surətdə 28332, 28523, 57312, 32309, 62566, 28229, 51751 və 21117 zülal ardıcıllığı, həmçinin 29 metabolik yol və ya funksional sistem üzrə 1770 (kəsəkotu) və 1823 (düyü) gen götürülmüşdür (cədvəl 1). Bu cür dəqiqləşdirmə («təmizləmə») əməliyyatı statistik cəhətdən daha dəqiq, düzgün nəticə əldə etmək məqsədi ilə aparılmışdır. Promotor analizlərində xloroplast təyinatlı nüvə genlərinin 5'-rayonlarının (-1020: +30; +1 ATG inisiasiya kodonuna uyğundur) nukleotid ardıcıllıqları istifadə olunmuşdur.

RNT polimeraza II promotorlarının (TSS-lərin) axtarışı **TSSP-TCM** kompüter proqramının (Shahmuradov et al., 2005) vasitəsi ilə həyata keçirilmişdir.

Zülalların lokalizasiya yerlərinin proqnozlaşdırılması **ProtComp** kompüter proqramının (http://linux1.softberry.com/berry.phtml) köməyi ilə aparılmışdır.

Müxtəlif növlərdən olan, ardıcıllıq və təyinat yeri baxımından oxşar zülal/gen qruplarının axtarışı **BLAST/BLAN** və **ProtComp** analizlərinin nəticələri əsasında **GetOrtholog** kompüter proqramı (İ.Şahmuradov; çap olunmamışdır) vasitəsi ilə təhlil edilmişdir.

NƏTİCƏLƏR VƏ ONLARIN MÜZAKİRƏSİ

8 bitki orqanizminin nüvə genomunda plastid təyinatlı nüvə genlərinin axtarışı

Əvvəlcə birləpəli və ikiləpəli bitkilərin 8 nümayəndəsinin annotasiya olunmuş nüvə genlərinin kodlaşdırdığı zülalların amin turşusu ardıcıllıqları ProtComp proqramının vasitəsi ilə analiz olunmuş onların hüceyrə daxilində və membranında potensial təyinat yerləri (daşındıqları ünvanlar) müəyyən edilmişdir. Bu araşdırmaya görə potensial təyinat yeri plastidlər olan genlər üzrə inteqrativ məlumat cədvəl 2-də təqdim olunur. 2 saylı cədvəldən göründüyü kimi, plastid təyinatlı nüvə genlərinin ümumi sayına görə müxtəlif növlər arasında böyük fərq alınır. Belə ki, qarğıdalı, qara yonca və soya bitkilərində plastid təyinatlı nüvə genlərinin sayı daha çoxdur (6500-8200), şərab üzümü isə bu göstəriciyə görə nisbətən kasad görünür. Martin və həmkarlarının (Martin et al., 1998; 2002) məlumatına əsasən, ali bitkilərin nüvə genomunda plastid ünvanlı zülal genlərinin sayı 4500-5000 intervalındadır.

Cədvəl 2. Birləpəli və ikiləpəli bitkilərin nüvə genomlarında plastid təyinatlı genlərin axtarışının inteqrativ nəticələri

Orqanizm	Plastid təyinatlı zülal kodlaşdırması güman olunan genlərin ümumi sayı
Arabidopsis thaliana	5121
Glycine max	7682
Medicago truncatula	6526
Oryza sativa japonica	3810
Populus trichocarpa	3621
Sorghum bicolor	4295
Vitis vinifera	3535
Zea mays	8222

Bu baxımdan bizim nəticələri necə şərh etmək olar? Güman edirik ki, müşahidə olunan təzadlı vəziyyət aşağıdakı bir neçə səbəbdən ola bilər.

(1) Müxtəlif bitkilərin nüvə genomlarının mövcud annotasiyası dəqiq deyildir: bəzi genlər orada qeyd olunmur və/ya səhv olaraq real gen kimi qeyd olunmuşlar.

(2) Zülalların təyinat ünvanları üzrə ProtComp proqramının əsaslandığı biliklər tam yaxud dəqiq deyildir və bu səbəbdən də əldə olunan proqnoz səciyyəli məlumatlarda müəyyən xəta mövcuddur.

(3) Həqiqətən də, müxtəlif növlərin plastid zülal dəstləri arasında müəyyən fərq vardır.

İstənilən halda, bu nəticələrin əlavə təcrübi və bioinformatik analiz vasitəsi ilə dəqiqləşdirilməsi tələb olunur.

BLAST/BLAN və **ProtComp** nəticələrinin **GetOrtholog** proqramının köməyi ilə sonrakı analizi müxtəlif növlərin təmsil olunduğu oxşar gen qruplarının axtarışı həyata keçirilmişdir. Bu araşdırmanın ümumi nəticələri Cədvəl 3-də verilir, bəzi nümunələri isə Cədvəl 4-də verilir. Bütöv zülal ardıcıllıqlarının 60% və daha yuxarı oxşarlıq dərəcəsi meyarı əsas götürülməklə, müxtəlif tip oxşar gen qrupları müəyyənləşdirilmişdir (o cümlədən, Cədvəl 3-də qeyd olunan tip qruplar). Cədvəldən göründüyü kimi, 4642 qrupda 3 birləpəli bitkinin hər biri, ən azı, bir gen ilə təmsil olunur. İkiləpəlilərdə 2042 belə qrup aşkar olunmuşdur. Müşahidə olunan fərq, ola bilər ki, daha çox müxtəlif ikiləpəli orqanizmin müqayisə olunması ilə bağlıdır. Aşkar olunan digər maraqlı və təəccüblü fakt ayrılıqda birləpəlilər və ikiləpəlilər üçün plastid təyinatlı ümumi ("şərikli") gen qruplarının nisbətən azlığıdır (müvafiq surətdə, 533 və 112). Bu nəticə müxtəlif səbəblərdən ola bilər. Məsələn, ola bilsin ki, bizim minimal oxşarlıq dərəcəsi kimi götürdüyümüz "≥60%" oxşarlıq həddi zülallar üçün çox yüksəkdir.

İstənilən halda, bu nəticələri də sonrakı araşdırmalar yolu ilə dəqiqləşdirilməlidir.

Nəhayət, burada bir məqam da vurğulanmalıdır ki, proqnozlaşdırılan oxşar gen qruplarının hər hansı birindəki genlərin/zülalların ortoloq olub-olmaması bu araşdırmalardan aydın deyildir - bunun üçün əlavə tədqiqatlar tələb olunur. Belə ki, ortoloq genlərin müəyyənləşdirilməsi daha mürəkkəb, illər tələb edən bir problemdir.

Kəsəkotunun və düyünün XTZK nüvə genlərində potensial promotorlarının axtarışı

Kəsəkotunda xloroplastların müxtəlif funksiyaları ilə bağlı 29 qrupa daxil olan zülallar kodlaşdıran 1770 nüvə geninin 5'-rayonlarının (-1020: +30) nukleotid ardıcıllıqlarında TSSP-TCM kompüter proqramının köməyi ilə statistik cəhətdən yüksək etibarlılıq dərəcəsinə malik potensial RNT polimeraza II promotorlarının axtarışı həyata keçirilmiş və 1159 (~66%) gen üçün potensial promotorlar aşkar olunmuşdur. Hər bir promotor üçün genin kodlaşdıran hissəsinin başlanğıcına (+1) nəzərən TSS (və ya TSS-lər) və hər bir TATA promotor üçün həm də müvafiq TATA-boks müəyyənləşdirilmişdir. 1159 genin əksəriyyəti (864; 74,5%) üçün yalnız bir potensial promotor tapılmışdır. Digər tərəfdən 295 (25,5%) gen üçün birdən çox (2-4; əksər hallarda 2 və yalnız bir gendə 4) promotor müəyyənləşdirilmişdir. Bu nəticələr inteqral formada Cədvəl 5-də verilmişdir.

Cədvəl 3. Birləpəli və ikiləpəli bitkilərin nüvə genomlarında oxşar gen qruplarının axtarışının inteqrativ nəticələri

Oxşar genlər qrupunun tipi	Tapılmış gen qruplarının ümumi sayı
Qrupda 4 və daha çox növ, ən azı, bir gen ilə təmsil olunur (yalnız müvafiq zülalların oxşarlığına görə - güman olunan təyinat yeri və bitkinin birləpəli yaxud ikiləpəli olması nəzərə alınmadan)	5072
Qrupda 3 birləpəli bitkinin hər biri, ən azı, bir gen ilə təmsil olunur (yalnız müvafiq zülalların oxşarlığına görə - güman olunan təyinat yeri nəzərə alınmadan)	4642
Qrupda 5 ikiləpəli bitkinin hər biri, ən azı, bir gen ilə təmsil olunur (yalnız müvafiq zülalların oxşarlığına görə - güman olunan təyinat yeri nəzərə alınmadan)	2042
Qrupda 4 və daha çox növ plastid təyinatlı zülal kodlaşdıran gen(lər) ilə təmsil olunur (yalnız müvafiq zülalların oxşarlığına görə - bitkinin birləpəli yaxud ikiləpəli olması nəzərə alınmadan)	657
Qrupda 3 birləpəli bitkinin hər biri plastid təyinatlı zülal kodlaşdıran bir və daha çox gen ilə təmsil olunur	533
Qrupda 5 ikiləpəli bitkinin hər biri plastid təyinatlı zülal kodlaşdıran bir və daha çox gen ilə təmsil olunur	112

Cədvəl 4. Plastid təyinatlı oxşar gen qruplarının nümunələri

Oxşar genlər qrupunun tipi	Qrupa daxil olan genlər
Qrupda 8 bitkinin hər biri, ən azı, plastid təyinatlı zülal kodlaşdıran bir gen ilə təmsil olunur	AT: NC_003071.7_cdsid_NP_179576.1 [gene=HXK2] AT: NC_003075.7_cdsid_NP_194642.1 [gene=HXK1] GM: Glyma01g01060.1|PACid:16242989 GM: Glyma05g35890.1|PACid:16260844 GM: Glyma07g12190.1|PACid:16267046 GM: Glyma08g03730.1|PACid:16269849 MT: IMGA|Medtr8g014530.1 OS: NC_008394.4_cdsid_NP_001044214.1 [gene=Os01g0742500] PT: NC_008484.1_cdsid_XP_002325031.1 [gene=POPTRDRAFT_825877] PT: NC_008467.1_cdsid_XP_002299739.1 [gene=POPTRDRAFT_815248] SB: Sb03g034230.1|PACid:1963247 SB: Sb09g026080.1|PACid:1981558 VV: NW_003724068.1_cdsid_XP_002283608.1 [gene=LOC100242358] VV: NW_003724055.1_cdsid_XP_002283574.1 [gene=LOC100244595] ZM: GRMZM2G171373_P01 ZM: GRMZM2G432801_P01 ZM: GRMZM2G058745_P01
Qrupda 3 birləpəli bitkinin hər biri plastid təyinatlı zülal kodlaşdıran bir və daha çox geni ilə təmsil olunur	OS: NC_008396.2_cdsid_NP_001051386.1 [gene=Os03g0767000] SB: Sb01g007000.1|PACid:1950143 ZM: GRMZM2G033098_P01 ZM: GRMZM2G376661_P01 ZM: GRMZM2G072653_P01
Qrupda 5 ikiləpəli bitkinin hər biri plastid təyinatlı zülal kodlaşdıran bir və daha çox geni ilə təmsil olunur	AT: NC_003071.7_cdsid_NP_178585.1 [gene=LHCB2.1] AT: NC_003071.7_cdsid_NP_178582.1 [gene=LHCB2.2] AT: NC_003074.8_cdsid_NP_189406.1 [gene=LHCB2.3] AT: NC_003070.9_cdsid_NP_564339.1 [gene=CAB3] AT: NC_003071.7_cdsid_NP_565786.1 [gene=LHB1B2] GM: Glyma02g47560.1|PACid:16250525 MT: IMGA|Medtr5g097280.1 MT: IMGA|contig_81211_1.1 MT: IMGA|contig_58796_1.1 MT: IMGA|Medtr6g012110.1 PT: NC_008480.1_cdsid_XP_002321186.1 [gene=Lhcb2-1] PT: NC_008468.1_cdsid_XP_002301582.1 [gene=Lhcb2-2] VV: NW_003724078.1_cdsid_XP_002283566.1 [gene=LOC100250504] VV: NW_003724082.1_cdsid_XP_002271687.1 [gene=LOC100254533] VV: NW_003724037.1_cdsid_XP_002273106.2 [gene=LOC100252066]

AT: *Arabidopsis thaliana*, GM: *Glycine max*, MT: *Medicago truncatula*, OS: *Oryza sativa japonica*, PT: *Populus trichocarpa*, SB: *Sorghum bicolor*, VV: *Vitis vinifera*, ZM: *Zea mays*.

Cədvəl 5. Kəsəkotu və düyünün XTZK nüvə genlərində potensial promotorların sayı

Bir gendə tapılmış promotorların ümumi sayı	Genlərin sayı (kəsəkotu)	Genlərin sayı (düyü)
≥1 promotor	1159	1491
1 promotor	864; 74,5%	906; 60,8%
2 promotor	258; 22.3%	493; 33,1%
3 promotor	36; 3,1%	83; 5,6%
4 promotor	1; 0,1%	8; 0,5%
5 promotor	0	1; 0,1%
>5 promotor	0	0

Cədvəl 6-da kəsəkotunun 29 qrup üzrə promotor axtarışının inteqral nəticələri verilmişdir. Müəyyən olunmuşdur ki, birdən çox promotor tapılmış genlər üçün, yalnız KH-nin başlanğıcına ən yaxın TSS götürülməsi şərti daxilində, potensial promotorlar içərisində TATA və qeyri-TATA tip promotorların nisbi payı baxımından qeyri-TATA tip promotorlar üstünlük (59%) təşkil edirlər. Bu üstünlük 8 qrup («Assemblinq faktorları», «Xloroplast membranları», «Translyasiya», «Birləşdirici zülallar», «Müdafiə mexanizmləri», «Elektron daşıyıcılarının sintezi və deqradasiyası», «Qocalma» və «Transkripsiya») üzrə daha qabarıqdır (boz rənglə qeyd olunmuşdur). Digər tərəfdən, 3 qrup («Stressə cavab reaksiyası», «İkincili metabolizm», «Sulfat metabolizmi») TATA promotorlar nisbətən çoxluq təşkil edir. Qalan qruplarda bu və ya digər tip promotorlar aşkar üstünlük təşkil etmirlər.

Düyüdə xloroplastların müxtəlif funksiyaları ilə bağlı 29 qrupa daxil olan zülallar kodlaşdıran 1823 nüvə geninin 5'-rayonlarının (-1020: +30) nukleotid ardıcıllıqlarında TSSP-TCM kompüter proqramının

Cədvəl 6. Kəsəkotunun XTZK nüvə genlərinin potensial promotorlarının axtarışının yekun nəticələri

Qruplar	Analiz olunan genlərin ümumi sayı	Heç olmasa, 1 (hər hansı tip) promotor tapılan genlər		Yalnız TATA promotor(lar) tapılan genlər	Yalnız qeyri-TATA promotor(lar) tapılan genlər	Hər 2 tip promotorlar tapılan genlər
		Cəmi	d<=100[1]			
Amin turşularının sintezi	181	118; 65,2%	53; 44,9%	38;32,2%	63;53,4%	17;14,4%
Assemblinq faktorları	8	6; 75%	3; 50%	1; 16,7%	5; 83,3%	0
Birləşdirici zülallar	10	6; 60%	2; 33,3%	2; 33,3%	4; 66,7%	0
Xloroplastın biogenezi	40	32; 80%	9; 28,1%	10; 31,2%	19; 59,4%	3; 9,4%
Karbohidrat metabolizmi	115	77; 67%	42; 54,5%	21; 27,3%	42; 54,5%	14; 18,2%
Şaperonlar və İştilik şoku zülalları	28	20; 71,4%	8; 40%	7; 35%	12; 60%	1; 5%
Xloroplastın inkişafı və difirensasiyası	43	24; 55,8%	9; 37,5%	8; 33,3%	14; 58,3%	2; 8,3%
Xloroplast membranı	28	13; 46,4%	9; 69,2%	1; 7,7%	11; 84,6%	1; 7.7%
Qaranlıq reaksiyası	66	40;75,8%	31; 62%	16; 32%	24; 48%	10; 20%
Müdafiə mexanizmləri	3	2; 66,7%	0	0	2; 100%	0
DNT-nin replikasiyası, reparasiyası və rekombinasiyası	28	16; 57,1%	7; 43,8%	4; 25,0%	10; 62,5%	2; 12,5%
Elektron daşıyıcılarının sintezi və deqradasiyası	2	1; 50%	0	0	1; 100%	0
Yağ turşuları metabolizmi	121	78; 64,5%	43; 55,1%	23; 29,5%	39; 50%	16; 20,5%
İşıq reaksiyası	76	50; 65,8%	33; 66%	14; 28%	29; 58%	7; 14;0%
İşıq reseptorlarının sintezi və deqradasiyası	7	5; 71,4%	3; 60%	2; 40%	3; 60%	0
Azot metabolizmi	148	95; 64,2%	41; 43,2%	29; 30,5%	46; 48,4%	20; 21,1%
Fosforlaşma və defosforlaşma faktorları	4	2; 50%	0	1; 50%	1; 50%	0
Piqmentlərin sintezi və deqradasiyası	57	39; 66,7%	17; 43,6%	10; 25,6%	22; 56,4%	7; 17,9%
Post-translyasiya modifikasiyası	27	11; 40,7%	6; 54,5%	3; 27,3%	5; 45,5%	3; 27,3%
Proteazalar	55	30; 54,5%	13; 43,3%	8; 26,7%	16; 53,3%	6; 20,0%
Stresə cavab reaksiyası ilə bağlı genlər	101	76; 75,2%	49; 64,5%	36; 47,4%	27; 35,5%	13; 17,1%
İkincili metabolizm	271	158; 58,3%	88; 55,7%	76; 48,1%	55; 34,8%	27; 17,1%
Qocalma	4	3; 75,0%	2; 66,7%	0	3; 100%	0
Sulfat metabolizm	53	35; 66,0%	16; 45,7%	18; 51,4%	13; 37,1%	4; 11,4%
Tilakoid membranlarınn translokon sistemi	11	6; 54,5%	3; 50%	2; 33,3%	2; 33,3%	2; 33,3%
Transkriptlərin prosessinqi	33	16; 48,5%	6; 37,5%	2; 12,5%	9; 56,2%	5; 31,2%
Transkripsiya	35	21; 60%	6; 28,6%	3; 14,3%	15; 71,4%	3; 14,3%
Translyasiya	99	60; 60,6%	26; 43,3%	9; 15%	42; 70%	9; 15%
Translokatorlar və transportyorlar	48	34; 70,8%	16; 47,1%	11; 32,4%	14; 41,2%	9; 26,5%
Bütün gruplar üzrə, CƏMİ	**1770**	**1159; 65,5%**	**569; 49,1%**	**387; 33,4%**	**569; 49,1%**	**203; 17,5%**

köməyi ilə statistik cəhətdən yüksək etibarlılıq dərəcəsinə malik potensial RNT polimeraza II promotorlarının axtarışı həyata keçirilmiş və 1491 (~82%) gen üçün potensial promotorlar aşkar olunmuşdur. Hər bir promotor üçün genin kodlaşdıran hissəsinin başlanğıcına (+1) nəzərən TSS (və ya TSS-lər) və hər bir TATA promotor üçün həm də müvafiq TATA-boks müəyyənləşdirilmişdir. 1491 genin əksəriyyəti (906; 60,8%) üçün yalnız bir potensial promotor tapılmışdır. Digər tərəfdən 585 (39,3%) gen üçün birdən çox (2-4; əksər hallarda 2) promotor müəyyənləşdirilmişdir (Cədvəl 5).

Cədvəl 7-də düyünün 29 qrup (1823 gen) üzrə promotor axtarışının inteqral nəticələri verilmişdir. Aşkar olunmuşdur ki, birdən çox promotor tapılmış genlər üçün yalnız KH-nin başlanğıcına ən yaxın TSS götürülməsi şərti daxilində, potensial promotorlar içərisində TATA və qeyri-TATA tip promotorların nisbi payı baxımından qeyri-TATA tip promotorlar mütləq üstünlük (74,4%) təşkil edirlər. Bu üstünlük 14 qrupda («Assemblinq faktorları», «Xloroplast membranları», «Translyasiya», «Amin turşularının sintezi» və b.) müşahidə olunur. TATA promotorlar nisbətən çoxluq təşkil edən yeganə gen qrup «Müdafiə mexanizmləri» qrupudur.

Kəsəkotu və düyüdə xloroplastların müxtəlif funksiyaları ilə bağlı 29 qrupa daxil olan zülallar kodlaşdıran nüvə genlərində RNT polimeraza II promotorlarının axtarışı ikiləpəli və birləpəli bitkilərin bu nümayəndələri arasında promotor tipi (TATA və ya qeyri-TATA) baxımından həm ümumi, həm də fərqli cəhətlər aşkar etmişdir.

Cədvəl 7. Düyünün XTZK nüvə genlərinin potensial promotorlarının axtarışının yekun nəticələri

Qruplar	Analiz olunan genlərin ümumi sayı	Heç olmasa 1 (hər hansı tip) promotor tapılan genlər		Yalnız TATA promotor(lar) tapılan genlər	Yalnız qeyri-TATA promotor(lar) tapılan genlər	Hər 2 tip promotorlar tapılan genlər
		Cəmi	d<=100			
Amin turşularının sintezi	174	145; 83,3%	71; 49%	27; 18,6%	88; 60,7%	30; 20,7%
Assemblinq faktorları	6	6; 100%	2; 33,3%	2; 33,3%	4; 66,7%	0
Birləşdirici zülallar	6	2; 33,3%	1; 50%	0	0	2; 100%
Xloroplastın biogenezi	42	35; 83,3%	15; 42,9%	5; 14,3%	24; 68,6%	6; 17,1%
Karbohidrat metabolizmi	137	109; 79,6%	59; 54,1%	15; 13,8%	62; 56,9%	32; 29,4%
Şaperonlar və İştilik şoku zülalları	47	38; 80,9%	18; 47,4%	7; 18,4%	24; 63,2%	7; 18,4%
Xloroplastın inkişafı və difirensasiyası	32	25; 78,1%	13; 52,0%	0	21; 84,0%	4; 16%
Xloroplast membranı	23	16; 69,6%	4; 25%	1; 6,2%	11; 68,8%	4; 25%
Qaranlıq reaksiyası	74	60; 81,1%	36; 60%	9; 15%	31; 51,7%	20; 33,3%
Müdafiə mexanizmləri	5	5; 100%	2; 40%	3; 60%	2; 40%	0
DNT-nin replikasiyası, reparasiyası və rekombinasiyası	30	24;80%	12; 50%	3; 12,5%	18; 75%	3; 12,5%
Elektron daşıyıcılarının sintezi və deqradasiyası	5	4; 80%	2; 50%	0	2; 50%	2; 50%
Yağ turşuları metabolizmi	143	119; 83,2%	50; 42%	21; 17,6%	72; 60,5%	26; 1,8%
İşıq reaksiyası	57	48; 84,2%	30; 62,5%	5; 10,4%	36; 75%	7; 14,6%
İşıq reseptorlarının sintezi və deqradasiyası	4	3; 75%	0	0	1; 33,3%	2; 66,7%
Azot metabolizmi	151	117; 77,5%	53; 45,3%	19; 16,2%	76; 65,0%	22;18,8%
Fosforlaşma və defosforlaşma faktorları	3	2; 66,7%	1; 50%	1; 50%	1; 50%	0
Piqmntlərin sintezi və deqradasiyası	52	43; 82,7%	22; 51,2%	5; 11,6%	26; 60,5%	12; 7,9%
Post-translyasiya modifikasiyası	24	18; 75%	8; 44,4%	6; 33,3%	11; 61,1%	1; 5,6%
Proteazalar	46	40; 87%	18; 45%	4; 10%	31; 77,5%	5; 12,5%
Stresə cavab reaksiyası ilə bağlı genlər	93	83; 89,2%	48; 57,8%	14; 16,9%	42; 50,6%	27; 2,5%
İkincili metabolizm	273	199; 72,9%	92; 46,2%	65; 32,7%	93; 46,7%	41; 0,6%
Qocalma	5	4; 80%	1; 25%	0	2; 50%	2; 50%
Sulfat metabolizmi	62	51; 82,3%	25; 49%	8; 15,7%	34; 66,7%	9; 17,6%
Tilakoid membranlarınn translokon sistemi	13	13; 100%	6; 46,2%	1; 7,7%	10; 76,9%	2; 15,4%
Transkriptlərin prosessinqi	33	32; 97%	21; 65,6%	2; 6,2%	23; 71,9	7; 21,9
Transkripsiya	23	21; 91,3%	9; 42,9%	3; 14,3%	12; 57,1%	6; 28,6%
Translyasiya	113	85; 75,2%	48; 56,5%	9; 10,6%	62; 72,9%	14; 6,5%
Translokatorlar və transportyorlar	75	57; 76%	22; 38,6%	10; 17,5%	34; 59,6%	13; 2,8%
Bütün gruplar üzrə, CƏMİ	**1823**	**1491; 81,8%**	**703; 47,1%**	**259; 17,4%**	**890; 59,7%**	**342; 22,9%**

<u>Ümumi cəhətlər:</u>

(1) bütövlükdə və qruplar üzrə əsasən xloroplast təyinatlı nüvə genlərində qeyri-TATA promotorlar üstünlük təşkil edir;

(2) hər iki orqanizm üzrə 3 eyni qrupda («Assemblinq faktorları», «Xloroplast membranları», «Translyasiya») qeyri-TATA tip promotorlar xüsusilə (2 dəfə və daha çox) üstünlük təşkil edirlər, 1 qrupda («Fosforlaşma və defosforlaşma faktorları») hər iki tip promotorlar təxminən eyni dərəcədə təmsil olunmuşdur.

<u>Fərqli cəhətlər:</u>

(1) 5 qrup («Birləşdirici zülallar», «Müdafiə mexanizmləri», «Elektron daşıyıcılarının sintezi və deqradasiyası», «Qocalma», «Transkripsiya») üzrə yalnız kəsəkotunda və 11 qrup («Amin turşularının sintezi», «Xloroplastın biogenezi», «Xloroplastın inkişafı və diferensiasiyası», «DNT-nin replikasiyası, reparasiyası və Rekombinasiyası», «İşıq reaksiyası», «Azot metabolizmi», «Piqmentlərin sintezi və deqradasiyası», «Sulfat metabolizmi», «Tilakoid membranlarınn translokon sistemi», «Transkriptlərin prosessinqi», «Proteazalar») üzrə yalnız düyüdə qeyri-TATA tip promotorlar xüsusilə (2 dəfə və daha çox) üstünlük təşkil edirlər.

(2) 3 qrup («Stressə cavab reaksiyası», «İkincil metabolizm», «Sulfat metabolizmi») üzrə yalnız kəsəkotunda və yalnız 1 qrup («Müdafiə mexanizmləri») üzrə düyüdə TATA promotorlar nisbi üstünlük təşkil edirlər. Maraqlıdır ki, düyüdə, bunun əksinə olaraq, «Stressə cavab reaksiyası» və «Sulfat metabolizmi» genlərində qeyri-TATA promotorlar üstünlük təşkil edirlər, «İkincili metabolizm» genlərində isə hər iki tip promotorlar təxminən eyni dərəcədə təmsil olunmuşlar.

Beləliklə, düyünün xloroplast təyinatlı nüvə genləri qeyri-TATA promotorları ilə daha zəngindirlər: TATA və qeyri-TATA tip promotorların nisbi payı kəsəkotunda 41%/59%, düyüdə isə 26%/74%-dir.

Bu analizlərin nəticələri prizmasında zülal kodlaşdıran bitki genlərində TATA və qeyri-TATA tip promotorların nisbəti ilə bağlı ədəbiyyatdan məlum olan iki faktı qeyd etmək maraqlı olardı.

Müxtəlif orqanizmlərdən götürülmüş 232 gendən yalnız 36-sında (~16%) promotor TATA tipdir. Analiz olunmuş 46 Rubisko, Rubisko aktivaza və *LHC* (işıq toplayıcı kompleks) genlərində isə yalnız 2 gendə promotor TATA tipdir. Digər tərəfdən, 1-ci fotosistemin komponentlərini kodlaşdıran 7 gendən hamısında promotor TATA tipdir (Nakamura et al., 2002). Bu faktlar bizim nəticələrlə əsasən uzlaşır.

Kəsəkotunun ekspressiya olunduğu təsdiq olunmuş 9653 nüvə genindən 4465-ində (46,3%) KH-nin başlanğıcına ən yaxın promotor TATA tip, 5188-ində (53,7%) isə KH-nin başlanğıcına ən yaxın promotor qeyri-TATA tipdir (Shahmuradov et al., 2005). Beləliklə, ən azı kəsəkotunun bütöv genomu səviyyəsində TATA və qeyri-TATA promotorların nisbi payları çox də kəskin fərqlənmirlər.

Bu müqayisələr göstərir ki, TATA və qeyri-TATA promotorların nisbi payları arasında bizim işdə, həmçinin Nakamura və həmkarlarının (Nakamura et al., 2002) analizlərində müşahidə olunan kəskin fərqin bütövlükdə birləpəli və ikiləpəli bitkilər üçün səciyyəvi olub-olmadığını aydınlaşdırmaq üçün əlavə tədqiqatlar tələb olunur. Lakin onu da qeyd etmək lazımdır ki, Nakamura və həmkarlarının analizləri ilə müqayisədə, bizim tədqiqatlarda təxminən 10 dəfə çox gen öyrənilmişdir, yəni bizim nəticələr statistik etibarlılıq baxımından müqayisə olunmaz dərəcədə daha inandırıcıdır.

ƏDƏBİYYAT

Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. (1997) Gapped BLAST and PSIBLAST: a new generation of protein database search programs. *Nucl. Acids Res.*, **25:** 3389-3402

Arabidopsis Genome Initiative (2000) Analysis if the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature,* **408:** 796-815.

Balaji J., Crouch J.H., Petite P.V., Hoisington D.A. (2006) A database of annotated tentative orthologs from crop abiotic stress transcripts. *Bioinformation,* **1:** 225-227.

Bryant N., Lloyd J., Sweeney C., Myouga F., Meinke D. (2011) Identification of nuclear genes encoding chloroplast-localized proteins required for embryo development in Arabidopsis. *Plant Physiology,* **155:** 1678-1689.

Ding J., Hu H. (2013) NIM, a novel computationnal method for predicting nuclear-encoded chloroplast proteins. *J. of Medical and Bioengineering,* **2:** 115-119.

Goff S., Ricke D., Lan T. et al. (2002) A draft sequence of the rice genome (*Oryza sativa L. ssp. japonica*). *Science,* **296:** 92-100.

Gruissem W., Tonkyn J. (1993) Control mechanisms of plastid gene expression. *Critical Reviews in Plant Sciences,* **12:** 19-55.

Hamilton J.P., Buell C.B. (2012) Advances in plant genome sequencing. *The Plant J.,* **70:** 177-190.

Heazlewood J.L., Tonti-Filippini J., Verboom R.E., Millar A.H. (2005) Combining experimental and predicted datasets for determination of the subcellular location of proteins in Arabidopsis. *Plant Physiol.*, **139:** 598-609.

Heazlewood J.L., Verboom R.E., Tonti-Filippini J., Small I., Millar A.H. (2007) SUBA: the Arabidopsis Subcellular Database. *Nucleic Acids Res.*, **35:** D213-D218.

Kim S., Kang J., Chung Y.J., Li J., Ryu K.H. (2008) Clustering orthologous proteins across phylogenetically distant species. *Proteins: Structure, Function, and Bioinformatics,* **71:** 1113-1122.

Lemon B., Tjian R. (2000) Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev.,* **14:** 2551-2569.

Mache R., Lerbs-Mache S. (2001) Chloroplast genetic system of higher plants: Chromosome replication, chloroplast division and elements of the transcriptional apparatus. *Current Science,* **80:** 217-224.

Martin, W., Stoebe B., Goremykin V., Hansmann S., Hasegawa M., Kowallik K. (1998) Gene transfer to the nucleus and the evolution of chloroplasts. *Nature,* **393:** 162-165.

Martin W., Rujan T., Richly E., Hansen A., Cornelsen S., Lins T., Leister D., Stoebe B., Hasegawa M., Penny D. (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci. USA,* **99:** 12246-12251.

Nakamura M., Tsunoda T., Obokata J. (2002) Photosynthesis nuclear genes generally lack TATA-boxes: a tobacco photosystem I gene responds to light through an initiator. *Plant J.,* **29:** 1-10.

Novina C.D., Roy A.L. (1996) Core promoters and transcriptional control. *Trends Genet.,* **12:** 351-355.

Ohler U., Liao G.C., Niemann H., Rubin G.M. (2002) Computational analysis of core promoters in the Drosophila genome. *Genome Biol.,* **3(12):** RESEARCH 0087.1-0087.12.

Pedersen A.G., Baldi P., Chauvin Y., Brunak S. (1999) The biology of eukaryotic promoter prediction – a review. *Computers & Chemistry,* **23:** 191-207.

Pfannschmidt T. (2003) Chloroplast redox signals: how photosynthesis controls its own genes. *Trends Plant Sci.,* **8:** 33-41.

Shahmuradov I.A., Solovyev V.V. (2003) PromH: promoters identification using orthologous genomic sequences. *Nucl. Acids. Res.,* **31:** 3540-3545.

Shahmuradov I.A., Solovyev V.V., Gammerman A.J. (2005) Plant promoter prediction with confidence estimation. *Nucleic Acids Res.,* **33:** 1069-1076.

Smale S.T. (1997) Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. *Biochimica et Biophysica Acta,* **1351:** 73-88.

Smale S.T. (2001) Core promoters: active contributors to combinatorial gene regulation. *Genes and Dev.,* **15:** 2503-2508.

Tanz S.K., Castleden I., Hooper C.M., Vacher M., Small I., Millar A.H. (2013) SUBA3: a database for integrating experimentation and prediction to define the SUBcellular location of proteins in Arabidopsis. *Nucleic Acids Res.,* **41:** D1185-1191.

Wang B., Du Q., Yang X., Zhang D. (2014) Identification and characterization of nuclear genes involved in photosynthesis in Populus. *BMC Plant Biology,* 14:81 (http://www.biomedcentral.com/1471-2229/14/81).

Wu F., Mueller L.A., Crouzillat D.C., Pétiard V., Tanksley S.D. (2006) Combining Bioinformatics and Phylogenetics to Identify Large Sets of Single-Copy Orthologous Genes (COSII) for Comparative, Evolutionary and Systematic Studies: A Test Case in the Euasterid Plant Clade. *Genetics,* **174:** 1407-1420.

Yu J., Hu S., Wang J., Wong G.K. et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science,* **296:** 79-92.

Zhao Y., Tang F., Cheng J., Li L., Xing G., Zhu Y., Zhang L., Wei H., He F. (2003) An initiator and its flanking elements function as a core promoter driving transcription of the Hepatopoietin gene. *FEBS Lett.,* **540:** 58-64.

Ядерные Гены Пластидного Назначения у Высших Растений

Х.Ф. Алиева[1], А.У. Абдулазимова[1], Н.Ш. Мустафаев[1], Л.М. Сулейманова[2], З.А. Абасзаде, И.А. Шахмурадов[1,2]

[1]*Институт ботаники НАНА*
[2]*Азербайджанский медицинский университет*

У 8 высших растений (однодольные *Oryza sativa*, *Sorghum bicolor* и *Zea mays*, двудольные *Arabidopsis thaliana*, *Medicago truncatula*, *Populus trichocarpa*, *Glycine max* и *Vitis vinifera*) осуществлен поиск ядерных генов, кодирующих белки (полипептиды) пластидного назначения, а также сравнительный анализ ядерных генов пластидного назначения арабидопсиса и риса по типу промотора (ТАТА или без-ТАТА). Было выявлено, что ядерный геном каждого из этих растений кодирует более 3000 белковых (полипептидных) генов, предположительно пластидного назначения. Между разными видами по числу таких ядерных генов наблюдается большая разница. На основании критерия сходства аминокислотной последовательности, величина которой составляла 60% и выше, были определены различные типы групп сходных генов. В 4642 группах, каждое из трех однодольных растений представлено, как минимум, одним геном. У двудольных были выявлены 2042 такие группы. У *A.thaliana* и *O.sativa*, для 1159 и 1494 ядерных генов хлоропластного назначения, связанных с 29 метаболическими путями или функциональными комплексами, выявлены потенциальные ТАТА и без-ТАТА промоторы. Среди этих потенциальных промоторов, по относительному числу ТАТА и без-ТАТА промоторов, у арабидопсиса и риса преобладают без-ТАТА промоторы (41%/59% у арабидопсиса, 26%/74% у риса).

Ключевые слова: *Высшее растение, пластида, ядро, геном, ген, белок, промотор, компьютерный анализ*

Plastid-Related Nuclear Genes of Higher Plants

**H.F. Aliyeva[1], A.U. Abdulazimova[1], N.Sh. Mustafayev[1],
L.M. Suleymanova[2], Z.A. Abaszade, I.A. Shahmuradov[1,2]**

[1]Institute of Botany,ANAS
[2]Azerbaijan Medical University

Search for nuclear genes encoding plastid proteins in 8 higher plants (monocot *Oryza sativa*, *Sorghum bicolor* and *Zea mays*, dicot *Arabidopsis thaliana*, *Medicago truncatula*, *Populus trichocarpa*, *Glycine max* and *Vitis vinifera*), as well as a comparative analisis of plastid-related nuclear genes of Arabidopsis and rice for the promoter type (TATA or TATA-less) have been performed. It was found that nuclear genomes of these organisms contain more than 3000 genes encoding protein (polypeptide) predicted to have the plastid destination. The significant difference between these species in total number of the plastid-related nuclear genes is observed. With 60% or higher similarity level of the full-length polypeptide sequences, various types of similar gene groups have been predicted. In 4642 groups, every monocot plant is represented, at least, by one gene. In dicots, 2042 such groups have been revealed in dicots. Putative TATA and/or TATA-less promoters for 1159 and 1491 chloroplast-targeted nuclear genes involved in 29 metabolic pathways or functional complexes of *A.thaliana* and *O.sativa*, respectively, have been identified. Comparison of the relative shares of TATA and TATA-less promoters in the genes of *Arabidopsis* (41%/59%) and rice (26%/74%) revealed that TATA-less promoters prevail on TATA promoters (mostly, in rice).

Key words: *Higher plant, plastid, nucleus, genome, gene, protein, promoter, computational analysis*